

REPLACED BY  
ART 34 AMST

## WHAT IS CLAIMED IS:

1. An isolated nucleic acid molecule, which encodes a fluorescent or chromo- protein, selected from the group consisting of:

(a) a nucleic acid which encodes a protein comprising the amino acid sequence as shown in SEQ ID NOs: 2, 4, 6, 8, 10, 12, 14, 16, 18, 20 or 22;

(b) a nucleic acid comprising a nucleotide sequence as shown in SEQ ID NOs: 1, 3, 5, 7, 9, 11, 13, 15, 17, 19 or 21;

(c) a nucleic acid that hybridizes under stringent conditions to the nucleic acid of (a) or (b) above;

(d) a nucleic acid that encodes a protein that has at least about 75% sequence identity to the amino acid sequence of (a) above;

(e) a nucleic acid that has at least about 70% sequence identity to the nucleotide sequence of (b) above;

(f) a nucleic acid which encodes a protein having at least one amino acid substitution, deletion or insertion in the amino acid sequence as shown in SEQ ID NOs: 2, 4, 6, 8, 10, 12, 14, 16, 18, 20 or 22.

(g) a derivative or mimetic of the nucleic acid of (a), (b), (c), (d), (e) or (f) above;

(h) a mutant of the nucleic acid of (a), (b), (c), (d), or (e) above;

(i) a nucleic acid which differs from the nucleic acid of (b), (c), (d), (e), (f), (g) or (h) above due to the degeneracy of genetic code; and

(j) a fragment of the nucleic acid of (a) or (b) above.

2. The nucleic acid molecule of claim 1, wherein said nucleic acid is isolated from an organism from a Class Hydrozoa.

3. The nucleic acid molecule of claim 1, wherein said nucleic acid is isolated from an organism from a Sub-order Anthomedusae

4. The nucleic acid molecule of claim 1, wherein said nucleic acid is isolated from a Genus *Phialidium*.

5. A vector comprising the nucleic acid molecule according to claim 1.

6. An expression cassette comprising (a) the nucleic acid molecule according to Claim 1; and (b) regulatory elements for the expression of said nucleic acid molecule in the desired host-cell.

7. A cell comprising the nucleic acid molecule according to claim 1, the vector according to claim 5, or the expression cassette according to claim 6.

8. A stable cell line comprising the nucleic acid molecule according to claim 1, the vector according to claim 5, or the expression cassette according to claim 6.

9. A transgenic plant comprising the nucleic acid molecule according to claim 1, the vector according to claim 5, or the expression cassette according to claim 6.

10. A transgenic animal comprising the nucleic acid molecule according to claim 1, the vector according to claim 5, or the expression cassette according to claim 6.

5 11. A method for producing a fluorescent or chromo- protein, said method comprising (a) providing a nucleic acid molecule according to claim 1 operably linked to suitable expression regulatory elements (b) expressing the fluorescent or chromo- protein from said nucleic acid molecule, and (c) isolating the protein substantially free of other proteins.

12. A nucleic acid molecule comprising a fragment of the nucleic acid molecule  
10 according to claim 1, said fragment encoding a peptide of at least 100 amino acids in length

13. A nucleic acid molecule having a sequence that is substantially the same as, or identical to a nucleotide sequence of at least 300 residues in length of the nucleic acid molecule according to claim 1.

14. An isolated fluorescent or chromo- protein selected from the group consisting of:

15 (a) a protein comprising the amino acid sequence as shown in SEQ ID NOs: 2, 4, 6, 8, 10, 12, 14, 16, 18, 20 or 22;

(b) a protein encoded by the nucleic acid molecule comprising a nucleotide sequence as shown in SEQ ID NOs: 1, 3, 5, 7, 9, 11, 13, 15, 17, 19 or 21;

20 (c) a protein that has at least about 75% sequence identity to the amino acid sequence of (a) or (b) above;

(d) a mutant of the protein of (a), (b) or (c) above;

(e) a protein having at least one amino acid substitution, deletion or insertion in the amino acid sequence as shown in SEQ ID NOs: 2, 4, 6, 8, 10, 12, 14, 16, 18, 20 or 22.

(f) a derivative of the protein of (a), (b), (c), (d) or (e) above;

25 (g) a fragment of the protein of (a), (b), (c), (d), (e) or (f) above; and

(h) a protein having a sequence that is substantially the same as, or identical to the amino acid sequence of at least 100 residues in length of (a) or (b) above.

15. A fusion protein comprising the protein according to claim 14.

16. An antibody specifically binding to the protein according to claim 14.

30 17. A kit comprising the nucleic acid according to claim 1, the vector according to claim 5, the expression cassette according to claim 6, the protein according to claim 14, the fusion protein according to claim 15, or a means for producing the same.

18. An oligonucleotide probe or primer comprising the nucleotide sequence capable of hybridizing to the nucleotide sequence selected from the group consisting of SEQ ID NOs. 1, 3,  
35 5, 7, 9, 11, 13, 15, 17, 19, 21.

19. A method for labeling a biological molecule, comprising coupling said biological molecule to the protein according to claim 14.

20. A method for labeling a cell comprising production of the protein according to claim 14 in the cell.

5        21. A method for labeling a cell organelle comprising production of the protein according to claim 14 fused to the suitable subcellular localization signal in the cell.

22. A method for analyzing a biological molecule, cell or cell organelle comprising detection of fluorescence signal from the protein according to claim 14 or 15.

10       23. A method for analyzing a biological molecule, cell or cell organelle comprising expression of the nucleic acid molecule according to claim 1 in a cell.

24. A method of detecting a biological molecule comprising detection of fluorescence signal from the protein according to claim 14 or 15.

Gly Tyr Gly Asp Ala Ser Val Gly Lys Val Asp Ala Gln Phe Ile Cys  
 35 40 45  
 Thr Thr Gly Asp Val Pro Val Pro Trp Ser Thr Leu Val Thr Thr Leu  
 50 55 60  
 Thr Tyr Gly Ala Gln Cys Phe Ala Lys Tyr Gly Pro Glu Leu Lys Asp  
 65 70 75 80  
 Phe Tyr Lys Ser Cys Met Pro Glu Gly Tyr Val Gln Glu Arg Thr Ile  
 85 90 95  
 Thr Phe Glu Gly Asp Gly Val Phe Lys Thr Arg Ala Glu Val Thr Phe  
 100 105 110  
 Glu Asn Gly Ser Val Tyr Asn Arg Val Lys Leu Asn Gly Gln Gly Phe  
 115 120 125  
 Lys Lys Asp Gly His Val Leu Gly Lys Asn Leu Glu Phe Asn Phe Thr  
 130 135 140  
 Pro His Cys Leu Tyr Ile Trp Gly Asp Gln Ala Asn His Gly Leu Lys  
 145 150 155 160  
 Ser Ala Phe Lys Ile Met His Glu Ile Thr Gly Ser Lys Glu Asp Phe  
 165 170 175  
 Ile Val Ala Asp His Thr Gln Met Asn Thr Pro Ile Gly Gly Gly Pro  
 180 185 190  
 Val His Val Pro Glu Tyr His His Ile Thr Tyr His Val Thr Leu Ser  
 195 200 205  
 Lys Asp Val Thr Asp His Arg Asp Asn Met Ser Leu Val Glu Thr Val  
 210 215 220  
 Arg Ala Val Asp Cys Arg Lys Thr Tyr Leu  
 225 230

&lt;210&gt; 3

&lt;211&gt; 705

&lt;212&gt; DNA

&lt;213&gt; Artificial sequence

&lt;220&gt;

&lt;223&gt; phiYFP-Y1 mutant of the phiYFP

&lt;400&gt; 3

atgcctagtg gagcactggt gttccacgga aagatcccat atgttggtga gatggaggga 60  
 aatggtgatg gacacacatt ctccattaga ggtaaagggt atggagatgc aagtgttggt 120  
 aaagttgatg cccaattcat ctgcacaact ggagatgtac cagttccatg gtcaacttta 180  
 gtaacaacac ttacttatgg tgcacaatgc ttcgccaaat atggtccaga attaaaggat 240  
 ttctacaaga gttgcatgcc tgaaggctat gtgcaggagc gtacaatcac atttgaaggg 300  
 gacggagtat ttaaaactcg cgctgaagtt acatttgaaa acggatctgt ttataaccga 360

gtcaaactta atggacaagg atttaagaaa gacggacatg tgcttggaag gaatcttgaa 420  
ttcaatttca cacctcattg tctttacatt tggggagatc aggctaataca tggtttgaag 480  
tctgctttca aaattatgca tgagattact ggatcaaaaag gagacttcat tgttgacagac 540  
cacacccaaa tgaacacacc cattgggtggt ggaccagtcc atgtccctga ataccatcat 600  
atgacatacc atgtcactct cagcaaagat gttactgatac acagggataa catgagcttg 660  
gttgaaaccg tacgggctgt ggattgcaga aaaacatatc tttaa 705

<210> 4

<211> 234

<212> PRT

<213> Artificial sequence

<220>

<223> phiYFP-Y1 mutant of the phiYFP

<400> 4

Met Pro Ser Gly Ala Leu Leu Phe His Gly Lys Ile Pro Tyr Val Val  
1 5 10 15  
Glu Met Glu Gly Asn Val Asp Gly His Thr Phe Ser Ile Arg Gly Lys  
20 25 30  
Gly Tyr Gly Asp Ala Ser Val Gly Lys Val Asp Ala Gln Phe Ile Cys  
35 40 45  
Thr Thr Gly Asp Val Pro Val Pro Trp Ser Thr Leu Val Thr Thr Leu  
50 55 60  
Thr Tyr Gly Ala Gln Cys Phe Ala Lys Tyr Gly Pro Glu Leu Lys Asp  
65 70 75 80  
Phe Tyr Lys Ser Cys Met Pro Glu Gly Tyr Val Gln Glu Arg Thr Ile  
85 90 95  
Thr Phe Glu Gly Asp Gly Val Phe Lys Thr Arg Ala Glu Val Thr Phe  
100 105 110  
Glu Asn Gly Ser Val Tyr Asn Arg Val Lys Leu Asn Gly Gln Gly Phe  
115 120 125  
Lys Lys Asp Gly His Val Leu Gly Lys Asn Leu Glu Phe Asn Phe Thr  
130 135 140  
Pro His Cys Leu Tyr Ile Trp Gly Asp Gln Ala Asn His Gly Leu Lys  
145 150 155 160  
Ser Ala Phe Lys Ile Met His Glu Ile Thr Gly Ser Lys Gly Asp Phe  
165 170 175  
Ile Val Ala Asp His Thr Gln Met Asn Thr Pro Ile Gly Gly Gly Pro  
180 185 190  
Val His Val Pro Glu Tyr His His Met Thr Tyr His Val Thr Leu Ser  
195 200 205

Lys Asp Val Thr Asp His Arg Asp Asn Met Ser Leu Val Glu Thr Val  
 210 215 220  
 Arg Ala Val Asp Cys Arg Lys Thr Tyr Leu  
 225 230

<210> 5  
 <211> 705  
 <212> DNA  
 <213> Artificial sequence  
 <220>  
 <223> phiYFP-M0 mutant of the phiYFP  
 <400> 5

atgcctagtg gagcactggt gttccacgga aagatcccat atgttggtga gatggaggga 60  
 aatgttgatg gacacacatt ctccattaga ggtaaagggt atggagatgc aagtgttggt 120  
 aaagttgatg cccaattcat ctgcacaact ggagatgtac cagttccatg gtcaacttta 180  
 gtaacaacac ttacttatgg tgcacaatgc ttcgccaaat atggtccaga attaaaggat 240  
 ttctacaaga gttgcatgcc tgaaggctat gtgcaggagc gtacaatcac atttgaaggg 300  
 gacggaaact ttaaaactcg cgctgaagtt acatttgaaa acggatctgt ttataaccga 360  
 gtcaaaactta atggacaagg atttaagaaa gacggacatg tgcttggaag gaatcttgaa 420  
 ttcaatttca cacctcattg tctttacatt tggggagatc aggctaataca tggtttgaag 480  
 tctgctttca aaattcgcca tgagattact ggatcaaaag gagacttcat tgttgcagac 540  
 cacacccaaa tgaacacacc cattggtggt ggaccagtc atgtccctga aaaccatcat 600  
 atgagctacc atgtcaagct cagcaaagat gttactgatc acagggataa catgagcttg 660  
 aaggaaaccg tacgggctgt ggattgcaga aaaacatatc tttaa 705

<210> 6  
 <211> 234  
 <212> PRT  
 <213> Artificial sequence  
 <220>  
 <223> phiYFP-M0 mutant of the phiYFP  
 <400> 6

Met Pro Ser Gly Ala Leu Leu Phe His Gly Lys Ile Pro Tyr Val Val  
 1 5 10 15  
 Glu Met Glu Gly Asn Val Asp Gly His Thr Phe Ser Ile Arg Gly Lys  
 20 25 30  
 Gly Tyr Gly Asp Ala Ser Val Gly Lys Val Asp Ala Gln Phe Ile Cys  
 35 40 45

phiYFP-M1 using mammalian-optimised codons (SEQ ID NOs: 09, 10, and 27). "Humanized" version of phiYFP-M1 was subjected for site directed and random mutagenesis to obtain green and cyan light emitting versions of the protein. Mutant fluorescent proteins with green and cyan fluorescence were obtained. The green mutant of the humanized phiYFP-M1, named phiYFP-M1G1, contained the following amino acid substitutions (as compared with phiYFP-M1): T65S, L148Q, Y203T, K231T, T232A (SEQ ID NOs: 17, 18, and 31). The cyan mutant of the humanized phiYFP-M1, named phiYFP-M1C1, contained the following amino acid substitutions (as compared with phiYFP-M1): L6Q, T65S, Y66W, N124K, C147Y, L148Q, Y203T, V224L (SEQ ID NOs: 19, 20, and 32). Excitation-emission spectra for this protein are shown at Figure 3A,B.

### Example 3

#### hydr1GFP cloning, sequencing and recombinant protein production

Bright green fluorescence was detected using a fluorescent microscope in a hydromedusa 1 (about 1 mm in length, Figure 4) of sub-order *Anthomedusae* (*Cnidaria*, *Hydrozoa*, *Anthomedusae*). To search for the gene responsible for the fluorescence in this jellyfish, a strategy based on screening of an expression cDNA library in *E. coli* was implemented. Amplified cDNA samples were prepared using a SMART cDNA amplification kit (Clontech) and cloned into the PCR-Script vector (Stratagene). About 10<sup>5</sup> recombinant clones were screened visually using a fluorescent stereomicroscope. Three fluorescent clones were identified, each encoding the same green fluorescent protein, which was named hydr1GFP. The nucleotide and amino acid sequences for this protein are shown in SEQ ID NOS: 11, 12, and 28. A comparison of hydr1GFP with *A. victoria* GFP is shown in Figure 1. hydr1GFP appears to be more similar to GFP (37% identity) than to fluorescent proteins from corals.

To facilitate protein purification, the coding region of hydr1GFP was cloned into pQE30 expressing vector (Qiagen), so that recombinant protein contained six-histidine tag at its N-terminus. After expression in *E. coli*, hydr1GFP was purified by the metal-affinity resin, TALON (Clontech). The excitation-emission spectra for hydr1GFP showed peaks at 474 nm and 494 nm (Figure 5). In contrast to wild type *A. victoria* GFP, the novel hydr1GFP protein possessed only one absorption-excitation peak, which may correspond to a deprotonated chromophore state.

### Example 4

#### hm2CP cloning, sequencing and recombinant protein production

Bright green fluorescence was detected in small hydromedusa 2 of sub-order *Anthomedusae* (*Cnidaria*, *Hydrozoa*, *Anthomedusae*, Figure 4) using fluorescent microscope. To search for FP from this jellyfish we chose a strategy based on screening of expression cDNA